

AD-A120 827

SIGNAL DETECTION: GAUSSIAN AND NON-GAUSSIAN MODELS WITH 1/1
NUISANCE PARAMETERS(U) SAN DIEGO STATE UNIV CA DEPT OF
MATHEMATICAL SCIENCES C B BELL JUL 82 3-82
N00014-80-C-0208

UNCLASSIFIED

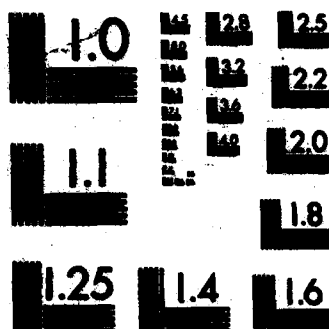
F/G 12/1

NL

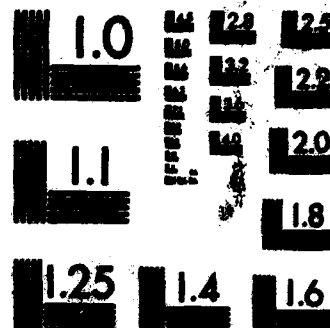
END

CHIEF

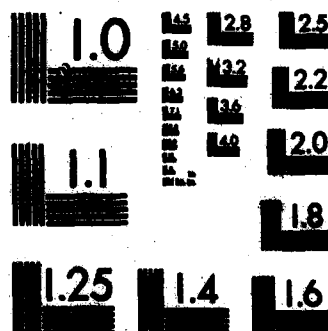
1/1



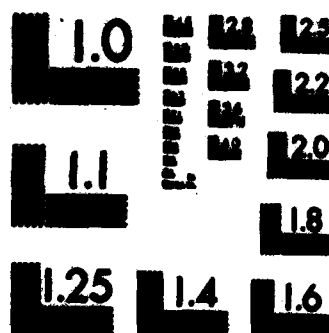
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



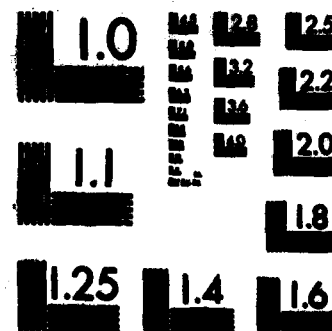
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ADA 120827

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 3-82	2. GOVT ACCESSION NO. AD A120827	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Signal Detection: Gaussian and Non-Gaussian Models with Nuisance Parameters		5. TYPE OF REPORT & PERIOD COVERED Technical
7. AUTHOR(s) C. B. BELL		6. PERFORMING ORG. REPORT NUMBER
8. PERFORMING ORGANIZATION NAME AND ADDRESS San Diego State University San Diego, CA 92182		9. CONTRACT OR GRANT NUMBER(s) N00014-80-C-0208
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE July 82
		13. NUMBER OF PAGES 33
		15. SECURITY CLASS. (of this report) Unclassified
		16. DECLASSIFICATION/DOWNGRADING SCHEDULE
18. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
19. SUPPLEMENTARY NOTES		
20. KEY WORDS (Continue on reverse side if necessary and identify by block number) Nuisance parameters, sufficient statistics, statistical noise, Poisson process, Weiner-Levy' process, uniform renewal process, randomized rank transformation, basic data transformation, Kolmogorov-, Lilliefors- and Srinivasan-type statistics.		
21. ABSTRACT (Continue on reverse side if necessary and identify by block number) Signal detection problems are treated in the context of several Gaussian and non-Gaussian stochastic processes with one or more nuisance parameters. The techniques all entail a basic data transformation and the use of a Kolmogorov-type statistic on one or more components of the transformed data. Extraneous statistical noise and randomized rank statistics are employed. Numerical examples are given.		

SIGNAL DETECTION: GAUSSIAN AND NON-GAUSSIAN

MODELS WITH NUISANCE PARAMETERS

C. B. Bell

San Diego State University

DTIC
SELECTED
S OCT 28 1982

A

1. Introduction and Summary.

One decision problem encountered in signal detection is the situation in which the law (L) of the process in the pure noise (PN) case is known precisely (i.e., $L = L^*$), and in the noise plus signal ($N + S$) case $L \neq L^*$. We assume the data consists of N observations X_1, \dots, X_N of the process. For this situation standard goodness-of-fit (G-O-F) detectors, such as those based on the Kolmogorov-Smirnov and Cramer-von Mises, statistics can be used.

One considers here extensions of the above problem to the case where the pure noise law is known up to some nuisance parameters. The detectors are all based on the maximal statistical noise (M-S-N) which is distributed independently of the minimal sufficient statistics (M-S-S) and together with the M-S-S constitute a one-to-one (a.e.) mapping of the data. This transformation is called the basic data transformation (BDT). The M-S-N and its relation for several common processes are presented in Section 2.

Lilliefors (1967, 1969) has developed procedures for the normal and exponential cases when nuisance parameters are present by using the

This work was supported by the Office of Naval Research through Grant No. N00014-80-C-0208.

This document has been approved
for public release and sale; its
distribution is unlimited.

maximum likelihood estimate of the distribution function. For the Kolmogorov-Smirnov statistic Srinivasan (1970) has developed similar procedures based on the Rao-Blackwell estimate of distribution function. Their results are presented in Section 3 along with extensions to the uniform distribution done by Choi (1980). Also, results for the normal case when the mean is known and the only nuisance parameter is the variance, are presented.

Section 4 illustrates how several other common detection procedures can be applied to the M-S-N when there are nuisance parameters. The types of procedures applied therein include (a) G-O-F (e.g., Cramer-von-Mises), (b) classical (e.g., F-statistic), and (c) nonparametric (e.g., sign-rank statistics).

The use of extraneous statistical noise (E-S-N) in detection procedures is introduced in Section 5. E-S-N is a simulated sample with known law L_0 belonging to the same family as L^* . By applying the inverse of the BDT one obtains, under PN, a sample with law L_0 . One can then apply standard detectors for this known distribution. The section concludes with a table of the inverse of the BDT for several stochastic process models of interest.

Finally, in Section 6 all of the above mentioned detection techniques are applied to computer generated data sets of five different processes. In addition, the techniques are applied to two real data sets consisting of interarrival times of epileptic seizures for two epileptic females.

2. Statistical Preliminaries: Maximal Statistical Noise (M-S-N).

Many decision procedures are based on the minimal sufficient statistic (M-S-S). This is particularly so when the inference is centered on the value of a parameter. However, in this paper one is primarily concerned with inference in the presence of nuisance parameters. Hence, the need to utilize something other than the M-S-S. This other quantity will be called the maximal statistical noise (M-S-N).

Definition 2.1 - Let $S(\underline{X})$ be the M-S-S for data $\underline{X} = (X_1, \dots, X_N)$, where the law, L , of the time series is an element of the family Ω' . Further, let $\delta(\underline{X}) = [N(\underline{X}), S(\underline{X})]$ be a transformation which is 1-1 a.e. and where $N(\underline{X})$ and $S(\underline{X})$ are statistically independent. Then

- (a) $N(\underline{X})$ is called M-S-N, i.e., maximal statistical noise; and
- (b) $\delta(\underline{X})$ is called the BDT, the basic data transformation.

Example 2.1 - (Homogeneous Poisson Process, HPP)

Let $[N(t): t \geq 0]$ be a HPP, i.e., homogeneous Poisson Process. Let (τ_n) and (W_n) be the associated interarrival time and waiting time series, respectively. If one views the data as $\underline{X} = (W_1, \dots, W_n)$, then one has

- (a) M-S-S: $S(\underline{X}) = W_N$
- (b) M-S-N $N(\underline{X}) = (Y_1, \dots, Y_{N-1}) = (\frac{W_1}{W_N}, \frac{W_2}{W_N}, \dots, \frac{W_{N-1}}{W_N})$ 1, and
- (c) BDT: $\delta(\underline{X}) = [N(\underline{X}), S(\underline{X})]$.

Example 2.2 - (Wiener-Levy Process, WLP)

Let $W(t) = \mu(t) + \sigma W^*(t)$, where $[W^*(t): t \geq 0]$ is a Wiener-Levy'

(WL) process satisfying (i) $E\{W^*(t)\} \equiv 0$; (ii) $\text{Cov}\{W^*(s), W^*(t)\} = \min(s, t)$; (iii) it is Gaussian. Let $Z_r = W(r\Delta)$, i.e., one samples at times $\Delta, 2\Delta, 3\Delta, \dots$, and $X = (X_1, \dots, X_N)$ where $X_r = Z_r - Z_{r-1}$ and $Z_0 = 0$.

Case I. $\mu(t) = \beta t$

Here, X_1, \dots, X_N are i.i.d. $N(\beta\Delta, \sigma^2\Delta)$;

$$\text{M-S-S: } S(X) = (\bar{X}, S_X) \text{ where } \bar{X} = N^{-1} \sum_{j=1}^N X_j, S_X^2 = N^{-1} \sum_{j=1}^N (X_j - \bar{X})^2$$

$$\text{M-S-N: } N(X) = \left(\frac{X_1 - \bar{X}}{S_X}, \dots, \frac{X_N - \bar{X}}{S_X} \right); \text{ and}$$

$$\text{BDT: } \delta(X) = [N(X), S(X)]$$

Case II. $\mu(t) \equiv 0$

Here, X_1, \dots, X_N are i.i.d. $N(0, \sigma^2\Delta)$;

$$\text{M-S-S: } S(X) = \sum_{j=1}^N X_j^2$$

$$\text{M-S-N: } N(X) = (e_1, \dots, e_N, V_1, \dots, V_N); \text{ where}$$

$$e_r = e(X_r) \text{ with } e(U) = 1 \text{ if } U \geq 0 \text{ and } = 0$$

$$\text{if } U < 0;$$

$$V_1 = \frac{X_2^2}{T_1}, V_2 = \frac{2X_3^2}{T_2}, V_3 = \frac{3X_4^2}{T_4}, \dots, V_{N-1} = \frac{(N-1)X_N^2}{T_N}$$

$$\text{and } T_r = \sum_{k=1}^r X_k^2.$$

[One should note here that if " σ " is a nondegenerate random variable, then $\{W(t)\}$ is non-Gaussian, but many of the detections techniques for Gaussian models still apply.]

Example 2.3 - (Uniform Renewal Process, URP)

Let U_1, \dots, U_N be i.i.d. $U(0, \theta)$ be the interarrival times of a (URP) point process. Then one has

$$M-S-S: S(U) = U(N);$$

$$M-S-N: N(U) = (R, V), \text{ where}$$

$$R = [R(U_1), \dots, R(U_N)] \text{ and } V = \left[\frac{U(1)}{U(N)}, \dots, \frac{U(N-1)}{U(N)} \right];$$

$$BDT: \delta(U) = [N(X), S(X)]$$

Example 2.4 - (Non-homogeneous Poisson Process, NHPP)

Let $[N^*(t): t \geq 0]$ be a NHPP whose mean function is known up to a multiplicative (nuisance) parameter α , i.e., $\mu^*(t) = \alpha \mu_0(t)$, and let $W^* = (W_1^*, \dots, W_N^*)$ be the waiting time data.

Now form $M(t) = N^*[\mu_0^{-1}(t)]$. Then $[M(t)]$ is a HPP with mean function $\mu(t) = \alpha t$ and waiting times $W = (W_1, \dots, W_N)$ where $W_r = \mu_0(W_r^*)$.

Hence, one has

$$M-S-S: S(W) = \mu_0(W_N^*);$$

$$M-S-N: N(W) = (V_1, \dots, V_{N-1}) \text{ with } V_r = \frac{\mu_0(W_r^*)}{\mu_0(W_N^*)};$$

$$BDT: \delta(W) = [N(W), S(W)].$$

Accession For

NAME GRA&I

NAME TAB

Unannounced

Classification

Letter on file

Classification/

Classification Codes



A

The statistics in the examples above involve the following distributions.

Definition 2.2 - (Basic Distributions)

- (1) If X_1, \dots, X_k are i.i.d $F(\cdot)$, continuous, one says

$$[R(X_1), \dots, R(X_k)] \sim R - V(k).$$

- (2) If U_1, \dots, U_k are i.i.d. $U(0,1)$ then

$$[U(1), \dots, U(k)] \sim U-O-S(k).$$

- (3) Let X_1, \dots, X_k be i.i.d. $N(\mu, \sigma^2)$, then

$$\left[\frac{X_1 - \bar{X}}{S_X}, \dots, \frac{X_k - \bar{X}}{S_X} \right] \sim N-M-S-N(k), \text{ where}$$

$$\bar{X} = \frac{1}{k} \sum_{i=1}^k X_i \text{ and } S_X^2 = \frac{1}{k} \sum_{i=1}^k (X_i - \bar{X})^2.$$

- (4) If $P(V < z) = z^k$, $0 < z < 1$, then one says $V \sim PW(k)$.

These processes and statistics are summarized in Table 2.1 below.

TABLE 2.1 PROCESSES AND STATISTICS

Stochastic Process	(Nuisance) Parameter	Data	M-S-S. (and Distn.) $S(X)$	M-S-N (and Distn.) $N(X)$
URP	θ	U_1, \dots, U_N	$\frac{U(N)}{\theta} \sim \text{PW}(N)$	(R, V) indep. with $R \sim R - V(N)$ and $V \sim U - 0 - S(N-1)$, where $R = [R(U_1), \dots, R(U_N)]$ $V_r = \frac{U(r)}{U(N)}, 1 < r < N-1$
HPP $\mu(t) = \lambda t$	λ	W_1, \dots, W_N	$W_N \sim \Gamma(n, \lambda)$ (i.e., the Gamma distribution with pars n , and λ .)	$\frac{W_1}{W_N}, \dots, \frac{W_{N-1}}{W_N} \sim \text{U-O-S}(N-1)$
MBPP $\mu(t) = \alpha \mu_0(t)$	α	W_1, \dots, W_N	$\mu_0(W_N) \sim \Gamma(n, 1)$ i.e., the Gamma distribution with pars $n, 1$.)	$\frac{\mu_0(W_1)}{\mu_0(W_N)}, \dots, \frac{\mu_0(W_{N-1})}{\mu_0(W_N)} \sim \text{U-O-S}(N-1)$
WLPL $\mu(t) = \beta t$	(β, σ^2)	X_1, \dots, X_N where $X_r = W(r\Delta) - W((r-1)\Delta)$ and $W(0) = 0$	(\bar{X}, S_X^2) indep. with $\bar{X} \sim N(\beta\Delta, \frac{\sigma^2\Delta}{n})$ $\frac{NS_X^2}{\sigma^2\Delta} \sim X_{N-1}^2$	$\frac{X_1 - \bar{X}}{S_X}, \dots, \frac{X_N - \bar{X}}{S_X} \sim \text{N-M-S-N}$
WLP $\mu(t) \equiv 0$	σ^2	Same as above	$\sum_{i=1}^N X_i^2 \sim \sigma^2 \Delta X_N$	(ϵ, V) indep.; (ϵ_r) i.i.d. $B(1, 1/2)$ $V_r \sim F(1, r), V_r$ indep., where $\epsilon_r = \epsilon(X_r), V_r = rX_{r+1}^2 (\sum_{i=1}^r X_i^2)^{-1}$

In order to implement the decision procedures based on the M-S-S. the construction utilizes GOF (goodness-of-fit) statistics and randomized tests.

GOF Procedures for Mixture Parameters

The GOF statistics to be used are the K-S (Kolmogorov-Smirnov) statistics and its various modifications in the presence of unknown mixture parameters. These are based on the empirical-type distribution function defined below.

Let X_1, X_2, \dots, X_n be a random sample of size n from a mixture distribution with

$$(i) \quad F(x) = \sum_{j=1}^k \pi_j F_j(x);$$

$$(ii) \quad f_j(x) = \frac{1}{\sigma_j} \phi\left(\frac{x - \mu_j}{\sigma_j}\right);$$

$$(iii) \quad \phi(z) = 1 - \exp(-z^2/2), \quad z > 0;$$

$$(iv) \quad G(\lambda; x) = 1 - \left(1 - \frac{x}{\lambda}\right)^{\lambda-1}, \quad 0 < x < \lambda;$$

$$(v) \quad G(n, \sigma; x) = \phi\left(\frac{x - \mu}{\sigma}\right);$$

$$(vi) \quad G(n, \sigma; x) = 1 - \int_0^x s(y) dy \quad \text{where} \quad v(x) = \frac{1}{2} \left(1 + \frac{x-2}{\sqrt{x}} \left(\frac{x}{x-1}\right)^{1/2}\right)$$

and $s(y) = \frac{\Gamma(x-2)}{[\Gamma(\frac{1}{2}(x-2))]^2} y^{1/2} x-2 (1-y)^{1/2} x-2, \quad 0 < y < 1;$

$$(viii) \quad G^*(\theta; z) = U(0, Y(n))$$

$$G^{**}(\theta; z) = \begin{cases} \left(\frac{k-1}{k}\right) \frac{z}{Y(n)}, & 0 < z < Y(n), \\ 1 & \text{when } z > Y(n). \end{cases}$$

Then,

$$(1) \quad D_1 = \sup_z |F_y^{(k)}(z) - G(z)| \sim K - S(k), \text{ i.e., has the K-S distribution}$$

for a random sample of size k (See Birnbaum, 1952).

$$(2) \quad D_2 = \sup_z |F_y^{(k)} - \hat{G}(\lambda; z)| \sim LLEX(k) \text{ if } G = \text{Exp}(\lambda) \text{ (See Lilliefors, 1970).}$$

$$(3) \quad D_3 = \sup_z |F_y^{(k)} - \tilde{G}(\lambda; z)| \sim SREX(k) \text{ if } G = \text{Exp}(\lambda) \text{ (See Srinivasan, 1970).}$$

$$(4) \quad D_4 = \sup_z |F_y^{(k)}(z) - \hat{G}(\mu; \sigma^2, z)| \sim LLGA(k) \text{ if } G = N(\mu, \sigma^2) \text{ (See Lilliefors, 1969).}$$

$$(5) \quad D_5 = \sup_z |F_y^{(k)}(z) - G(\mu, \sigma^2; z)| \sim SRGA(k) \text{ if } G = N(\mu, \sigma^2) \text{ (See Srinivasan, 1970).}$$

$$(6) \quad D_6 = \sup_z |F_y^{(k)}(z) - G^*(\theta; z)| \sim CHL(k), \text{ if } G = U(0, \theta) \text{ (See Choi, 1980 and Table A.1).}$$

$$(7) \quad D_7 = \sup_z |F_y^{(k)}(z) - G^{**}(\theta; z)| \sim CHS(k), \text{ if } G = U(0, \theta) \text{ (See Choi, 1980 in which it is proved that } D_7 \sim \left(\frac{k-1}{k}\right)K - S(k-1).)$$

The last six statistics are useful when the form of G is known, but the value of the (nuisance) parameter is unknown. These are based on the works of Lilliefors (1967, 1969), Srinivasan (1970), and Choi (1980).

who give the relevant tables. S. M. Lee computed an improved version of Choi's original table. It is Table A.1 of the appendix.

The utilization of these statistics for the detection models of this paper can be summarized in the table below, where usage of randomized noise is indicated. This concept will be defined in the next section.

Table 3.1. GOF Statistics

<u>Stochastic Process</u>	<u>GOF Statistics</u>	<u>Use of Randomized Noise</u>
URP	CHL, CHS, K-S	Yes, with K-S
WLPL $\mu(t) \equiv \mu t$	LLGA, SRGA	Yes, with K-S
WLP $\mu(t) \equiv 0$	LILE, SRLE (See discussion below)	Yes, with K-S
HPP $\mu(t) \equiv \lambda t$	LLEX, SREX, K-S	Yes, with K-S
NHPP $\mu(t) = \alpha \mu_0(t)$	LLEX, SREX, K-S	Yes, with K-S

The GOF statistic for the WLP with $\mu(t) \equiv 0$ has not been studied by the above named authors. However, one can derive the appropriate statistic in a manner parallel to those of Lilliefors, Srinivasan and Choi.

The Lilliefors idea is to replace the unknown parameter by its MLE (maximum likelihood estimate) in the distribution function. Srinivasan's approach is quite different but sometimes asymptotically equivalent. His idea is to replace the empirical distribution function $F_y^{(k)}(\cdot)$ by its Tao-Blackwell estimate. (See e.g., Bickel and Doksum, 1977).

Definition 3.2.

Let (1) $\hat{G}(\sigma; z) = \Phi(z\sqrt{k} [\sum_1^k y^2]^{-1/2})$; and

$$\tilde{G}(\sigma; z) = \begin{cases} 1 - 1/2 F(k-1, 1; (\sum_1^k x_r^2 - z^2)(N-1)^{-1} z^{-2}) & \text{for } 0 < z < (\sum_1^k x_r^2)^{1/2} \\ 1 & \text{if } z > (\sum_1^k x_r^2)^{1/2} \end{cases}$$

with $\tilde{G}(\sigma; z) + \tilde{G}(\sigma; -z) = 1$ for all z .

Here $F(m, r; \cdot)$ is the cdf of Fisher's F-statistic with m and r degrees of freedom.

Further define

$$(3) D_8 = \sup_z |F_y^{(k)}(z) - \hat{G}(\sigma; z)| \sim \text{LILE}(k) \text{ if}$$

$$G = N(0, \sigma^2); \text{ and}$$

$$(4) D_9 = \sup_z |F_y^{(k)}(z) - \tilde{G}(\sigma; z)| \sim \text{SRLE}(k) \text{ if}$$

$$G = N(0, \sigma^2)$$

Tables for these two statistics are given in the appendix.

It can be proved that the K-S statistic is a function of the M-S-S when the parameters are completely specified. The other eight statistics above satisfy a different property.

Proposition 3.1 - The GOF statistics in (2) - (7) of Def. 3.1 and (3) - (4) of Def. 3.2 are functions of the data solely through their respective M-S-N's (The proof will only be given in 2 cases.)

Proof for $\hat{G}(\mu, \sigma^2; z)$

$$\sup_z \left| \frac{1}{k} \sum_{i=1}^k \varepsilon(z - Y_i) - \Phi\left(\frac{z - \bar{Y}}{S_y}\right) \right| = \sup_w \left| \frac{1}{k} \sum_{i=1}^k \varepsilon(w - V_i) - \Phi(w) \right|,$$

$$\text{where } V_i = \frac{Y_i - \bar{Y}}{S_y}$$

Proof for $\check{G}(\lambda; z)$

$$\begin{aligned} \sup_{0 < z < k\bar{Y}} \left| \frac{1}{k} \sum_{i=1}^k \varepsilon(z - Y_i) - 1 + \left(1 - \frac{z}{k\bar{Y}}\right)^{k-1} \right| \\ = \sup_{0 < U < 1} \left| 1 - (1 - U)^{k-1} - \frac{1}{k} \sum_{i=1}^k (U - V_i + V_{i-1}) \right| \\ \text{where } V_i = \frac{\sum_{j=1}^i Y_j}{k\bar{Y}}. \end{aligned}$$

4. Generalized Randomized Rank Procedures

In order to avoid certain problems with distributions of statistics of interest, Durbin (1961), Bell and Doksum (1965) and others introduced extraneous noise into the decision procedures.

Example 4.1 - Let $Z = (Z_1, \dots, Z_N)$ be i.i.d. with continuous cdf $F(\cdot)$. Let $\xi = (\xi_1, \dots, \xi_N)$ be i.i.d. Φ and independent of the data Z . Now define $\xi' = (\xi'_1, \dots, \xi'_N)$ where ξ' is a permutation of ξ satisfying $\xi'_k = \xi(R(Z_k))$, i.e.,
$$\sum_{r=1}^N \varepsilon(\xi'_k - \xi'_r) = \sum_{r=1}^N \varepsilon(Z_k - Z_r).$$
 Bell and Doksum (1965) then prove that $\xi' \stackrel{d}{=} \xi$. Consequently,

$$\frac{1}{m} \sum_{j=1}^m \xi(R(Z_j)) - \frac{1}{N-m} \sum_{j=1}^N \xi(R(Z_j)) \sim N(0, \frac{1}{m} + \frac{1}{N-m}).$$

In order to formalize and generalize this procedure, one needs the following definitions.

Definition 4.1 - Let (r_1, \dots, r_N) be some permutation of the integers $1, \dots, N$, and let $v = (v_1, \dots, v_N)$ be an N -vector. τ_N^* is called the randomized rank transformation (RRT), when
$$\tau_N^*(r_1, \dots, r_N, v_1, \dots, v_N) = \{v(r_1), \dots, v(r_N)\} = v', \text{ i.e., } v'$$
 is a permutation of v such that
$$\sum_{s=1}^N \varepsilon(v(r_k) - v_s) = r_k \text{ for } 1 \leq k \leq N.$$

Besides the RRT defined above, the procedure in Ex. 4.1 involves an interchange of MSS's. For that example the MSS's are the order statistics.

Definition 4.2 - Let $X = (X_1, \dots, X_N)$ and $\xi = (\xi_1, \dots, \xi_N)$ be independent initial segments of time series with laws L and L^* , respectively, both in the family Ω' of distributions. Let $\delta(X) = \{N(X), S(X)\}$ be the BDT and $\xi' = \delta^{-1}\{N(X), S(\xi)\}$. Then, if X is the original data,

- (1) ξ_{\sim} is called extraneous statistical noise, ESN, and
 - (2) ξ'_{\sim} is called randomized statistical noise, RSN.
- (Note: L^* is chosen to be convenient and tractable.)

The principal result in this direction follows from

Lemma 4.1 - Let $Y = (Y_1, \dots, Y_N)$ have law $L \in \Omega'$; and

- (a) $N_{\sim 1} \stackrel{d}{=} N(Y)$;
- (b) $S_{\sim 1} \stackrel{d}{=} S(Y)$. Then $\delta^{-1}(N_{\sim 1}, S_{\sim 1}) \stackrel{d}{=} Y$.

Adapting the work of Durbin (1961), one has the example below.

Example 4.2 - (WLP with $\mu(t) = \mu t$)

Let $X_{\sim} = (X_1, \dots, X_N)$ be as in Ex. 2.2 and let $\xi_{\sim} = (\xi_1, \dots, \xi_N)$ be i.i.d ϕ and independent of X . One has then that

- (1) $\delta(X) = (\frac{X_1 - \bar{X}}{S_X}, \dots, \frac{X_N - \bar{X}}{S_X}, \bar{X}, S_X)$; and
- (2) $\delta^{-1}(\frac{X_1 - \bar{X}}{S_X}, \dots, \frac{X_N - \bar{X}}{S_X}, \bar{X}, S_X) = \xi'_{\sim} = (\xi'_1, \dots, \xi'_N)$, where

$$\xi'_r = \bar{\xi} + S_{\xi} \left(\frac{X_r - \bar{X}}{S_X} \right). \text{ Further,}$$

- (3) $\xi'_{\sim} \stackrel{d}{=} \xi_{\sim}$, and $\sup_z |F_{\xi}^{(N)}(z) - \phi(z)| \sim K - S(N)$.

An example for non-homogeneous Poisson processes is as follows:

Example 4.3 - (NHPP with $\mu(t) = \alpha(t^2 + t)$)

As in Ex. 2.4, let $\underline{W} = (W_1, \dots, W_N)$ be the waiting times. Then

$$S(\underline{W}) = (W_N^2 + W_N); \quad N(\underline{W}) = \left(\frac{W_1^2 + W_1}{W_N^2 + W_N}, \dots, \frac{W_{N-1}^2 + W_{N-1}}{W_N^2 + W_N} \right). \quad \text{Now, choose}$$

$\underline{\xi} = (\xi_1, \dots, \xi_N)$ to be the waiting time of a HPP with $\lambda = 2.5$, and independent of \underline{W} .

Further, let $\underline{\xi}' = \delta^{-1}(N(\underline{W}), S(\underline{\xi}))$, i.e., $\xi'_T = (N(\underline{\xi})) \left(\frac{W_T^2 + W_T}{W_N^2 + W_N} \right)$.

Then $\underline{\xi}' \stackrel{d}{=} \underline{\xi}$; and $\xi'_m (\xi'_N - \xi'_m)^{-1} \left(\frac{N-m}{m} \right) \sim F(2m, 2N-2m)$.

One completes this section with a table of δ^{-1} for the relevant stochastic processes.

Table 4.1 - Implementation of the Randomized Noise Theorem: The Inverse BDT, δ^{-1}

Stochastic Process	Nuisance Parameters	$v = \delta(x_N)$ (See Table 2.1)	$u = \delta^{-1}(y_N)$
URP	θ	$(x_1, \dots, x_N; v_1, \dots, v_N)$	$\tau_N^*(x_1, \dots, x_N, v_1 v_N, v_2 v_N, \dots, v_{N+1} v_N, v_N)$
HPP $\mu(t) = \lambda t$	λ	(v_1, \dots, v_N)	$(v_1 v_N, v_2 v_N, \dots, v_{N-1} v_N, v_N)$
MHP $\mu(t) = \alpha u_0(t)$	α	(v_1, \dots, v_N)	$(v_1 v_N, v_2 v_N, \dots, v_{N-1} v_N, v_N)$
MLP $\mu(t) \equiv \mu t$	(μ, σ)	$(v_1, \dots, v_N, v_{N+1}, v_{N+2})$	$(v_1 v_{N+2} + v_{N+1}, \dots, v_N v_{N+2} + v_{N+1})$
MLP $\mu(t) = 0$	σ	$(\epsilon_1, \dots, \epsilon_N, v_1, \dots, v_{N-1}, v_N)$	$[(2\epsilon_1 - 1) \sqrt{w_1}, \dots, (2\epsilon_N - 1) \sqrt{w_N}]$ <p>where $w_1 = (N-1)! v_N^{D-1}$, $w_2 = (N-1)! v_N v_1^{D-1}$ $w_r = D^{-1} [(r-1)!]^{-1} [(N-1)!] v_N^{r-1}$ $r-2 \sum_{\alpha=1}^{k-1} (v_j + j), 3 \leq r \leq N; D = \sum_{j=1}^N (v_j + j)$</p>

6. Numerical Examples

To illustrate the use of the detectors listed in the previous sections, data sets from five stochastic processes were computer generated. The five processes are:

- (1) HPP with rate parameter $\lambda = 5$.
- (2) NHPP with mean function $\mu(t) = \alpha\mu_0(t)$
where $\mu_0(t) = t^2 + t$ and $\alpha = 5$.
- (3) WLP ϕ , the Wiener-Levy' process with mean function
 $\mu(t) \equiv 0$ and $\sigma^2 = 5$, $\Delta = 2$.
- (4) WLPL, the Wiener-Levy' process with mean function
 $\mu(t) = \beta t$ where $\beta = 5$, $\sigma^2 = 5$ and $\Delta = 2$.
- (5) URP with $\theta = 5$.

In each case the number of observations is 20. The waiting times for the five data sets are listed and plotted in Figures 6.1 to 6.5.

Each data set was used with its respective pure noise type of detectors as well as with one alternative group of detectors. The false alarm rate (FAR) was set at .01 for all the tests. The results are shown in Tables 6.1 to 6.5.

In addition to the simulated data, two real data sets consisting of interarrival times of epileptic seizures was obtained from Choi (1980). The first data set, denoted Epilepsy I, contains 30 observations on an eight year old epileptic female recorded from 9:00 AM to 9:00 PM. The other data set, Epilepsy II, contains 20 observations on a twelve

year old epileptic female recorded from 7:02 AM to 7:02 PM. The waiting times are listed and plotted in Figures 6.1 to 6.7 of the appendix.

All 5 sets of detectors were applied to the Epilepsy I data and the HPP and NHPP detectors were applied to the Epilepsy II data with FAR = .01. The results are shown in Tables 6.1 to 6.5 of the appendix.

REFERENCES

1. Basawa, Ishwar V., and Rao, B. L. S. Prakasa (1980). "Statistical Inference for Stochastic Processes", Academic Press, New York.
2. Bell, C. B. (1964). Automatic Distribution-Free Statistical Signal Detection, NEL Report No. 1245, U.S. Navy Electronics Laboratory, San Diego, CA.
3. Bell, C. B. and Doksum, K. A. (1965). Some new distribution free statistics. Am. Math. Stat. 36, 203-214.
4. Bell, C. B., Woodroffe, M. F., and Avadhani, T. V. (1970). "Non-parametric tests for stochastic processes" in Nonparametric Techniques in Statistical Inference, Ed. by Madan Lal Puri, Cambridge Univ. Press.
5. Bell, C. B. (1973). "Statistical inference for some special stochastic processes" in Statistical Inference and Related Topics, ed., by Madan Lal Puri, Academic Press, Inc. 273-290.
6. Bell, C. B., Ramirez, F., and Smith, Eric (1980). Wiener-Levy Models, Spherically Exchangeable Time Series, and Simultaneous Inference in Growth Curve Analysis," Chapter 5 in "Advanced Asymptotic Testing and Estimation: A Symposium in honor of Wassily Hoeffding" ed. by I. M. Chakrovarti, Cambridge University Press.
7. Birnbaum, Z. W. (1952). Numerical tabulation of the distribution of Kolmogorov's statistic for finite sample size, J. Am. Statist. Ass. 47, 425-441.

8. Choi, Youn Jung (1980). Kolmogorov-Smirnov Test with Nuisance Parameters in the Uniform Case. Master of Science Thesis, University of Washington.
9. Durbin, J. (1961). Some methods of constructing exact tests. *Biometrika* 48, 41-55.
10. Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown, *J. Am. Statist. Ass.* 62, 399-402.
11. Lilliefors, H. W. (1969). On the Kolmogorov-Smirnov test for exponential distribution with mean unknown, *J. Am. Statist. Ass.* 64, 387-389.
12. Srinivasan, R. (1970). An approach to testing the goodness-of-fit of incompletely specified distributions, *Biometrika* 57 (3), 605-611.

APPENDIX

Graphs, Tables and Numerical Examples.

Table 6.1 PN: HPP N+S: Not HPP (FAR=.01)

Data	Simulated HPP (N=20)			Simulated URP (N=20)			Epilepsy I (Choi, 1980) (N=30)			Epilepsy II (Choi, 1980) (N=20)		
	Statistic Value	Critical Value	Decision	Statistic Value	Critical Value	Decision	Statistic Value	Critical Value	Decision	Statistic Value	Critical Value	Decision
K-S(M-S-N)	.2349	.3612	PN	.1766	.3612	PN	.2878	.2947	PN	.2113	.3612	PN
C-VN(M-S-N)	.1314	.7435	PN	.1254	.7435	PN	.7770	.7435	N+S	.2395	.7435	PN
F(N,N)(M-S-N)	1.139	.34, 2.94	PN	1.6490	.34, 2.94	PN	.50	.42, 2.39	PN	.5633	.34, 2.94	PN
K-S(E-S-N)	.2097	.3524	PN	.1143	.3524	PN	.2880	.2899	PN	.4608	.3521	N+S
C-VN(E-S-N)	.3097	.7435	PN	.0566	.7435	PN	.9105	.7435	N+S	1.4554	.7435	N+S
F(N,N)(E-S-N)	1.1441	.34, 2.94	PN	1.6974	.34, 2.94	PN	.50	.42, 2.39	PN	.5633	.34, 2.94	PN
LLEX	.1313	.278	PN	.1870	.278	PN	.1100	.226	PN	.3063	.278	N+S
SREX	.1399	.24	PN	.1782	.24	PN	.1168	.20	PN	.3155	.24	N+S

TABLE 6.2. PN: NHPP with $\mu(t) = \alpha\mu_0(t)$ N + S: NHPP with $\mu(t) \neq \alpha\mu_0(t)$ (FAR = .01)

Data Vector Statistic	Simulated NHPP (N = 20)			Simulated HPP (N = 20)			EPILEPSY I (Choi 1980)			EPILEPSY II (Choi 1980)		
	Statistic Value	Critical Value	Decision	Statistic Value	Critical Value	Decision	Statistic Value	Critical Value	Decision	Statistic Value	Critical Value	Decision
E-S (M-S-M)	.1634	.3612	PN	.2540	.3612	PN	.5276	.2947	N+S	.4548	.3612	N+S
C-W (M-S-M)	.0882	.7435	PN	.2722	.7435	PN	3.5304	.7435	N+S	1.2202	.7435	N+S
Fisher-Pearson (M-S-M)	31.4441	61.2	PN	56.1463	61.2	PN	157.1938	86.0	N+S	64.1642	61.2	N+S
T(M,N) (M-S-M)	1.1834	.34, 2.94	PN	.4857	.34, 2.94	PN	.1256	.42, 2.39	PN	.1496	.34, 2.94	N+S
E-S (E-S-M)	.1261	.3624	PN	.6510	.3624	N+S	.9900	.2899	N+S	.9928	.3524	N+S
C-W (E-S-M)	.0883	.7435	PN	3.2779	.7435	N+S	9.8998	.7435	N+S	6.6318	.7435	N+S
T(M,N) (E-S-M)	1.5340	.34, 2.94	PN	1.1441	.34, 2.94	PN	.4894	.42, 2.39	PN	.5639	.34, 2.04	PN
LLEX	.1678	.278	PN	.1769	.278	PN	.2350	.226	N+S	.3140	.278	N+S
SSEX	.1564	.24	PN	.1823	.24	PN	.2393	.20	N+S	.3230	.24	N+S

TABLE 6.4. PN: WLPL N + S: Not WLPL (FAR = 0.01)

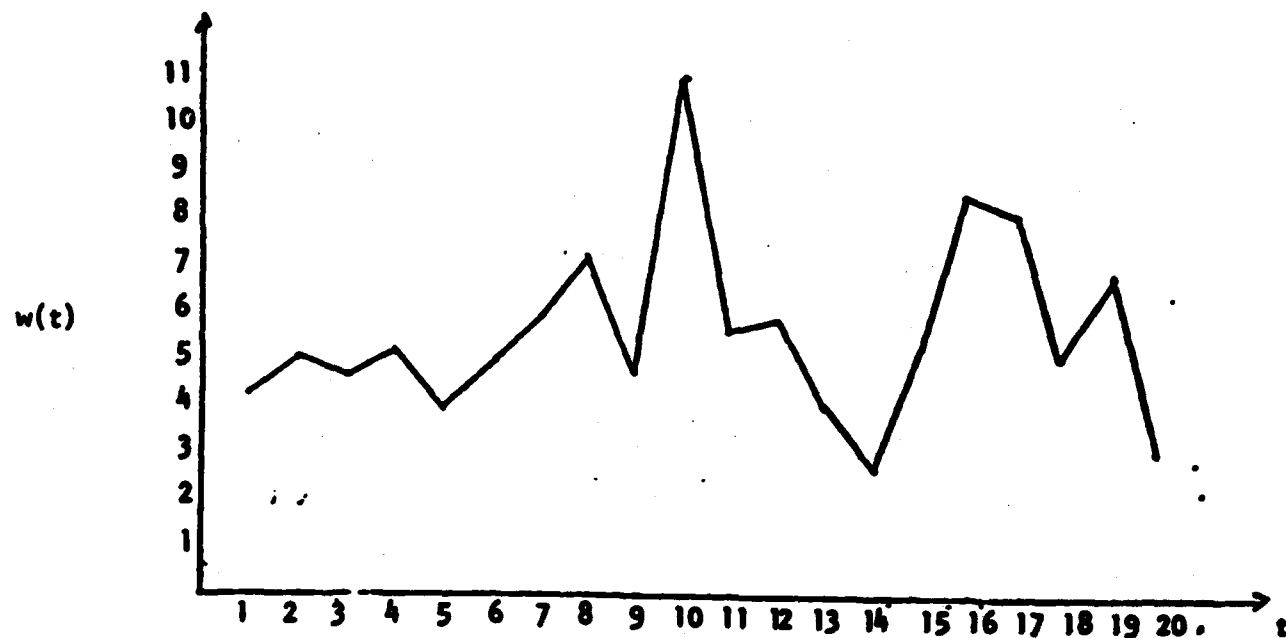
Data Detector Statistic	Simulated WLPL (N = 20)		Simulated WLPL (N = 20)		EPIDSI 1 (Choi 1980)	
	Statistic Value	Critical Value	Decision	Statistic Value	Critical Value	Decision
E-S (E-S-N)	.2063	.3524	PN	.1441	.3524	PN
$t(\frac{N}{2})$ (E-S-N)	.8807	+3.169	PN	.9808	+3.169	PN
LGA	.1380	.231	PN	.0867	.231	PN
				.2223	.2899	PN
				.8773	+2.602	PN
				.2241	.187	N+S

TABLE 6.5. PN: URP N + S: Not URP (FAR = .01)

Data Detector Statistic	Simulated WLPL (N = 20)		Simulated NRP (N = 20)		EPIDPSY I (Choi 1980)	
	Statistic Value	Critical Value	Statistic Value	Critical Value	Statistic Value	Critical Value
E-S (H-S-H)	.1683	.3612	.5367	.3612	.6742	.2047
C-VH (H-S-H)	.0567	.7435	2.1677	.7435	3.5484	.7435
E-S (E-S-H)	.2287	.3624	.4869	.3624	.5641	.2690
C-VH (E-S-H)	.3031	.7435	1.8463	.7435	3.4382	.7435
CHL	.1463	.3606	.4066	.3606	.5455	.3000
CHS	.1676	.3431	.5116	.3431	.5561	.2646

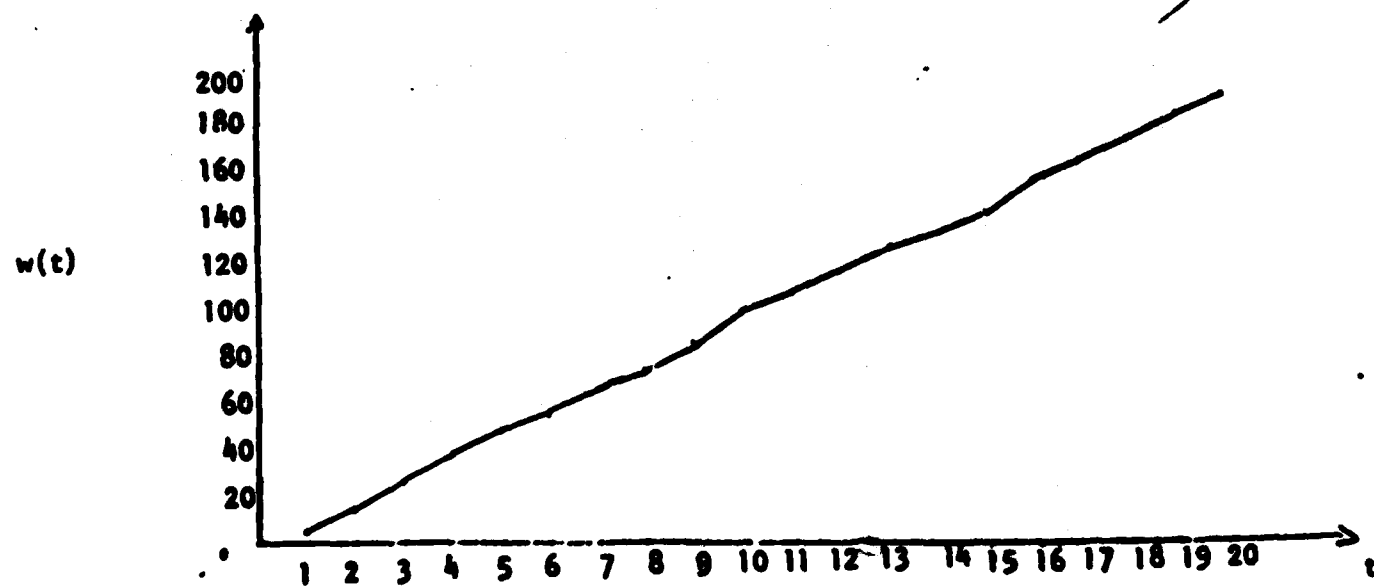
Waiting Times for WLPB Data

Figure 6.1.

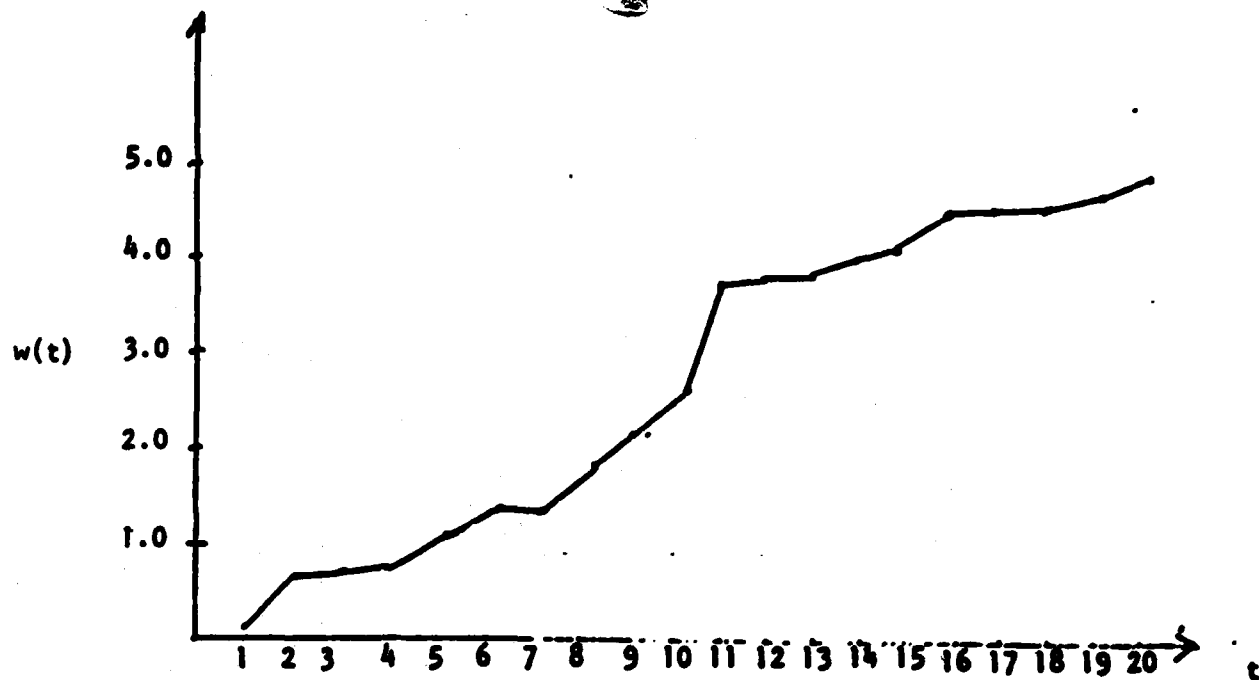


Waiting Times for WLPL Data

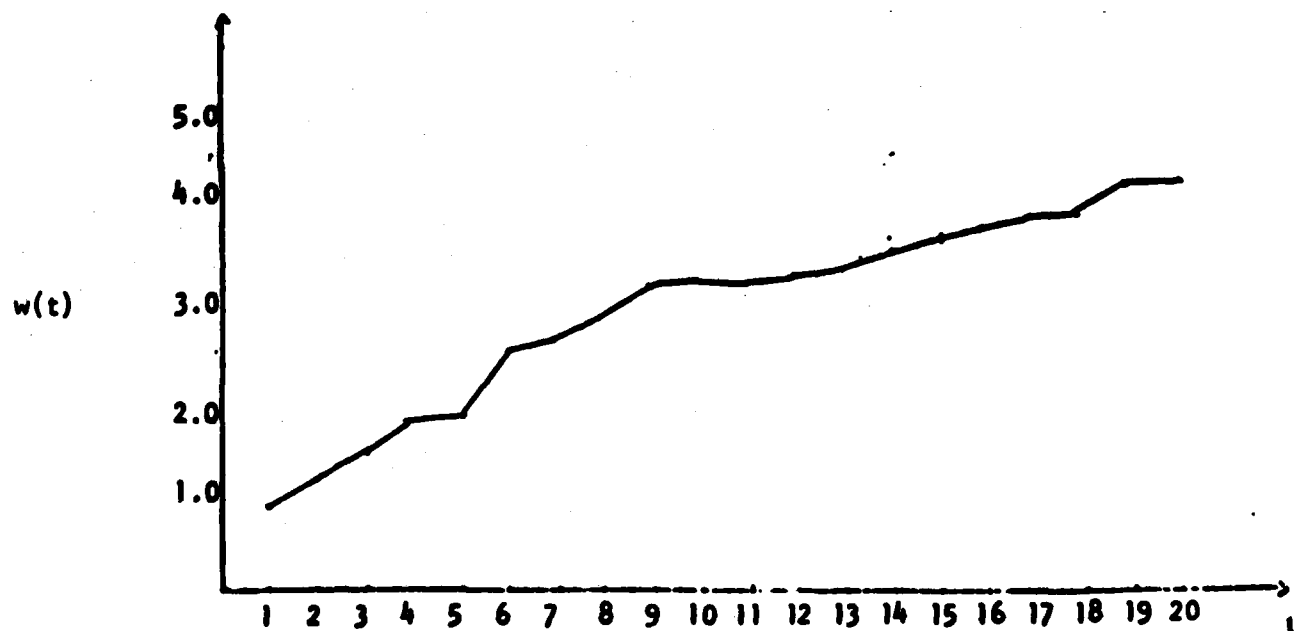
Figure 6.2.



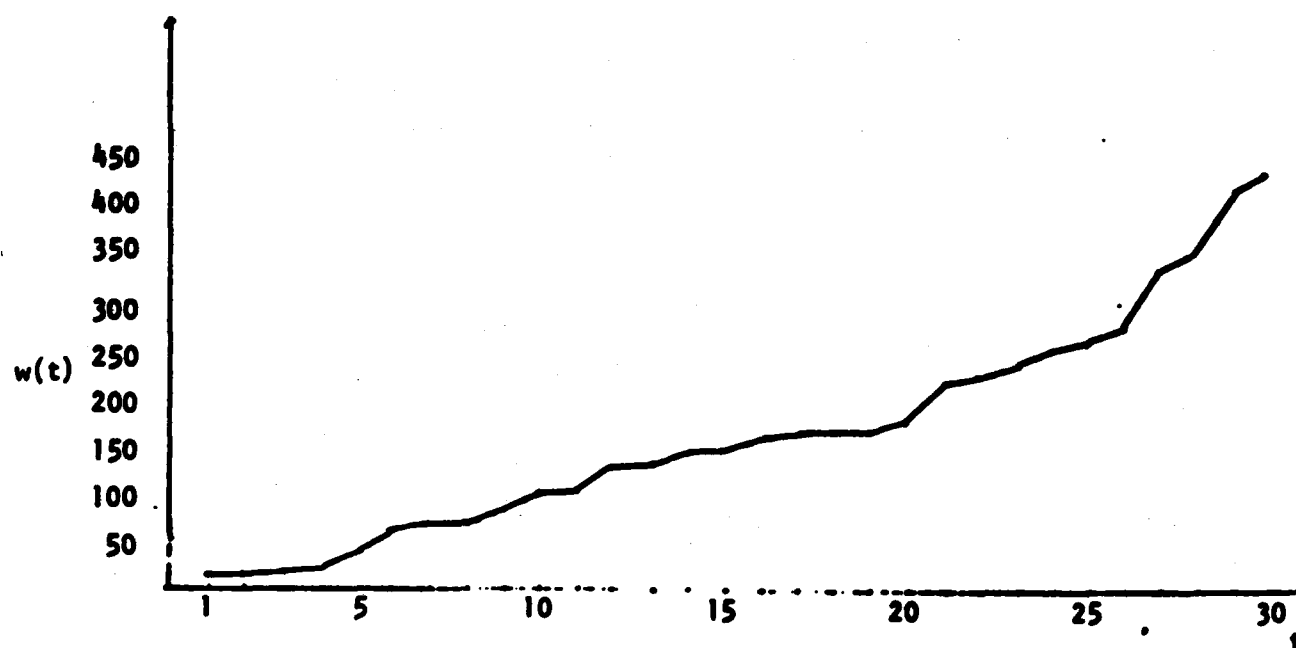
Waiting Times for HPP Data Figure 6.3.



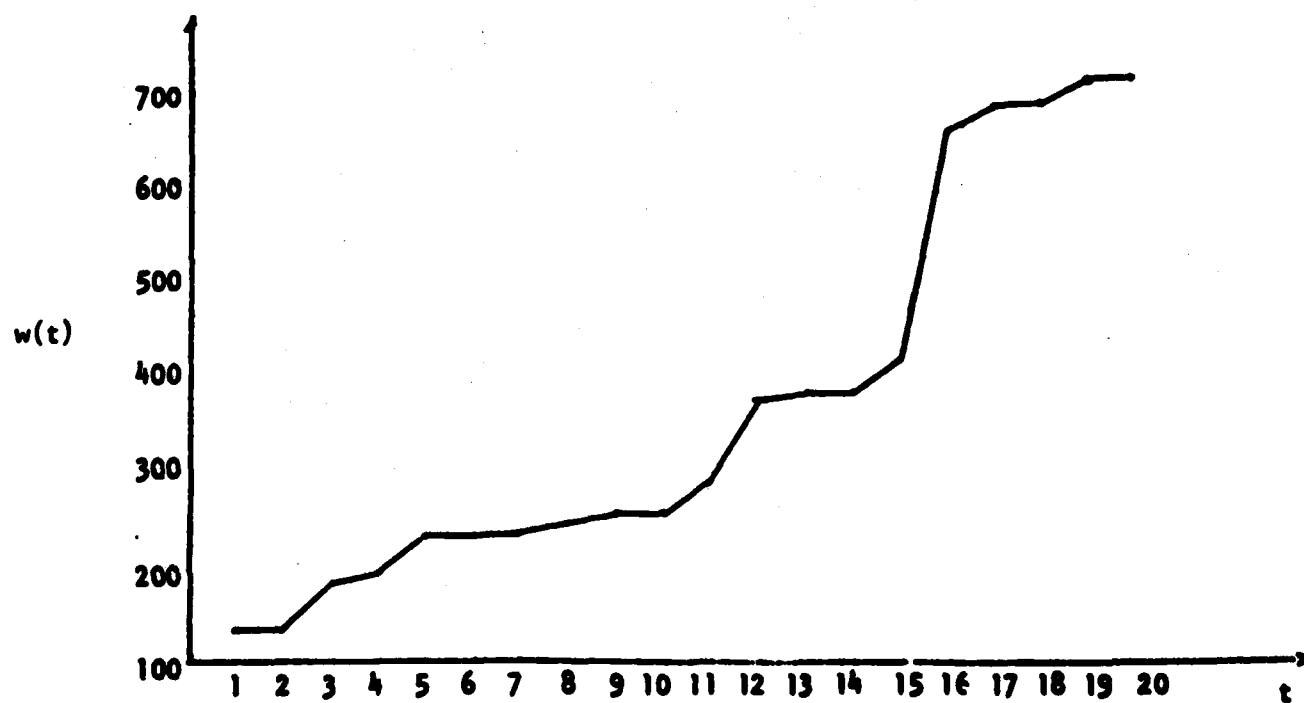
Waiting Times for NHPP Data Figure 6.4.



Waiting Time for Epilepsy I Data (secs.) Figure 6.5.



Waiting Times for Epilepsy II Data (secs.) Figure 6.6.



Waiting Times for URP Data Figure 6.7.

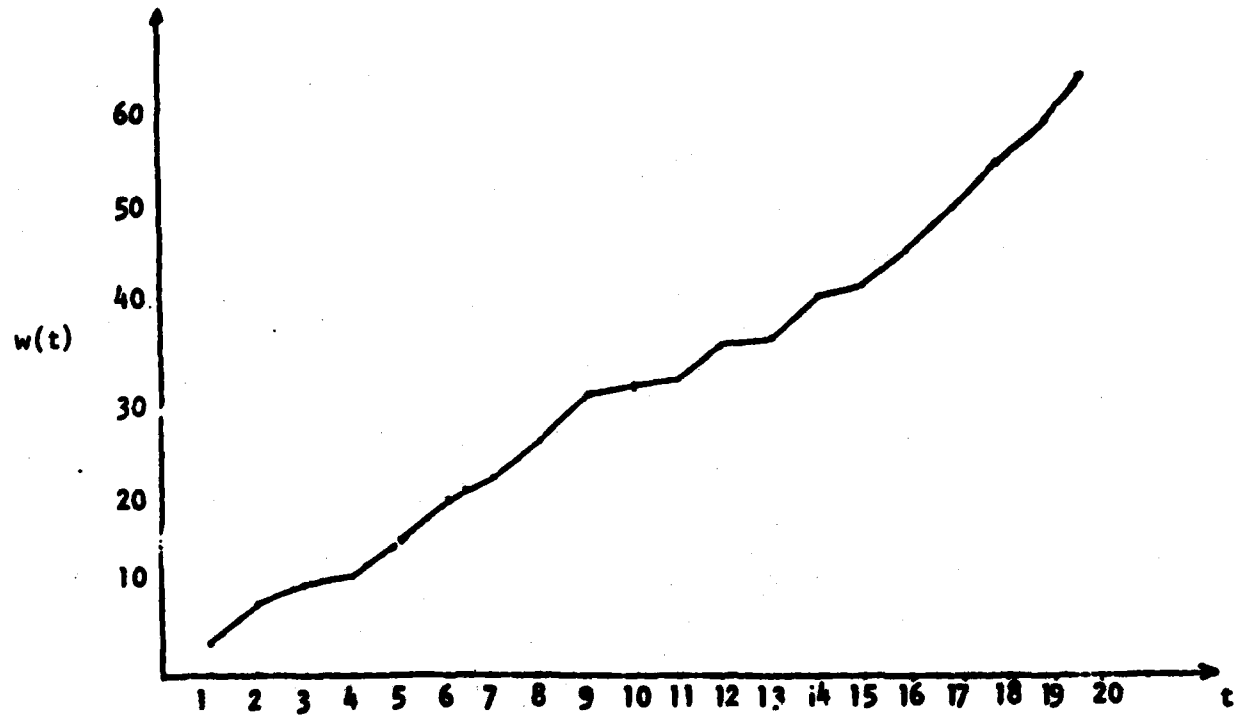


Table A.1 Critical Values of $D_{\alpha, CHL(k)}$ (See Section 3)

α k	.01	.05	.10	.15	.20	.25	.30
2	0.9909	0.9507	0.8996	0.8501	0.8001	0.7509	0.7018
3	0.8963	0.7746	0.6846	0.6472	0.6222	0.5967	0.5718
4	0.7878	0.6686	0.6071	0.5613	0.5241	0.4950	0.4783
5	0.7120	0.5935	0.5407	0.5030	0.4723	0.4465	0.4229
6	0.6482	0.5463	0.4922	0.4569	0.4304	0.4065	0.4229
7	0.5991	0.5032	0.4541	0.4216	0.3975	0.3779	0.3608
8	0.5603	0.4681	0.4224	0.3917	0.3685	0.3494	0.3341
9	0.5324	0.4457	0.3984	0.3699	0.3480	0.3290	0.3144
10	0.5025	0.4218	0.3789	0.3519	0.3313	0.3145	0.2993
11	0.4765	0.4009	0.3612	0.3342	0.3153	0.2987	0.2854
12	0.4618	0.3839	0.3473	0.3235	0.3033	0.2874	0.2742
13	0.4436	0.3724	0.3332	0.3084	0.2906	0.2761	0.2631
14	0.4263	0.3562	0.3203	0.2971	0.2788	0.2644	0.2523
15	0.4103	0.3419	0.3086	0.2853	0.2698	0.2555	0.2437
16	0.4012	0.3347	0.3019	0.2799	0.2636	0.2490	0.2470
17	0.3862	0.3245	0.2917	0.2713	0.2555	0.2421	0.2301
18	0.3813	0.3159	0.2823	0.2630	0.2476	0.2350	0.2245
19	0.3669	0.3057	0.2752	0.2555	0.2403	0.2276	0.2169
20	0.3553	0.2968	0.2672	0.2473	0.2332	0.2214	0.2113
21	0.3493	0.2918	0.2636	0.2445	0.2303	0.2179	0.2075
22	0.3401	0.2835	0.2567	0.2386	0.2245	0.2128	0.2030
23	0.3338	0.2789	0.2498	0.2317	0.2179	0.2069	0.1974
24	0.3244	0.2708	0.2430	0.2260	0.2127	0.2025	0.1933
25	0.3250	0.2671	0.2407	0.2232	0.2101	0.1995	0.1899
26	0.3146	0.2616	0.2350	0.2184	0.2060	0.1954	0.1864
27	0.3060	0.2564	0.2313	0.2156	0.2031	0.1925	0.1838
28	0.3029	0.2539	0.2286	0.2118	0.1996	0.1898	0.1811
29	0.2989	0.2481	0.2233	0.2073	0.1952	0.1849	0.1763
30	0.2914	0.2427	0.2187	0.2027	0.1908	0.1813	0.1732

(Table computed by S. M. Lee)

Table A.2 Critical Values of $D_{g,LILE}(k)$ (See Section 3)

$k \backslash \alpha$.01	.05	.10	.15	.20	.25	.30
2	0.8369	0.8202	0.7968	0.7697	0.7393	0.7062	0.6700
3	0.7975	0.7217	0.6566	0.5992	0.5507	0.5294	0.5107
4	0.7335	0.6210	0.5650	0.5289	0.4985	0.4701	0.4426
5	0.6603	0.5665	0.5074	0.4695	0.4403	0.4182	0.3969
6	0.6071	0.5140	0.4639	0.4289	0.4002	0.3790	0.3615
7	0.5663	0.4766	0.4288	0.3994	0.3752	0.3534	0.3355
8	0.5337	0.4497	0.4048	0.3739	0.3511	0.3322	0.3159
9	0.5048	0.4250	0.3823	0.3530	0.3318	0.3137	0.2981
10	0.4772	0.4005	0.3609	0.3351	0.3145	0.2978	0.2825
11	0.4659	0.3866	0.3458	0.3201	0.2999	0.2836	0.2695
12	0.4438	0.3709	0.3318	0.3069	0.2886	0.2725	0.2586
13	0.4278	0.3577	0.3198	0.2964	0.2774	0.2622	0.2491
14	0.4095	0.3411	0.3067	0.2838	0.2656	0.2513	0.2387
15	0.3973	0.3310	0.2974	0.2753	0.2589	0.2450	0.2328
16	0.3872	0.3215	0.2875	0.2674	0.2513	0.2374	0.2261
17	0.3769	0.3130	0.2819	0.2599	0.2439	0.2301	0.2186
18	0.3690	0.3024	0.2712	0.2511	0.2356	0.2226	0.2115
19	0.3567	0.2966	0.2655	0.2455	0.2311	0.2183	0.2078
20	0.3479	0.2896	0.2598	0.2408	0.2260	0.2137	0.2029
21	0.3369	0.2820	0.2521	0.2333	0.2193	0.2075	0.1973
22	0.3344	0.2781	0.2473	0.2284	0.2141	0.2023	0.1927
23	0.3189	0.2704	0.2415	0.2233	0.2097	0.1986	0.1888
24	0.3172	0.2627	0.2361	0.2183	0.2048	0.1938	0.1847
25	0.3082	0.2558	0.2306	0.2139	0.2015	0.1906	0.1809
26	0.3054	0.2551	0.2266	0.2095	0.1968	0.1863	0.1774
27	0.3016	0.2500	0.2244	0.2076	0.1943	0.1838	0.1750
28	0.2903	0.2444	0.2189	0.2031	0.1906	0.1803	0.1718
29	0.2897	0.2416	0.2169	0.2001	0.1879	0.1782	0.1689
30	0.2843	0.2356	0.2121	0.1975	0.1853	0.1748	0.1662

(Table computed by S. M. Lee; 20,000 repetitions)

Table A.3 Critical Values of D_9 , $SRLE(k)$ (See Section 3)

$k \backslash \alpha$.01	.05	.10	.15	.20	.25	.30	.35	.40
2	0.7451	0.7176	0.6907	0.6632	0.6380	0.6145	0.5917	0.5734	0.5484
3	0.6633	0.6104	0.5747	0.5435	0.5229	0.5019	0.4774	0.4591	0.4441
4	0.6025	0.5500	0.5203	0.4955	0.4767	0.4608	0.4398	0.4252	0.4117
5	0.5497	0.5123	0.4791	0.4590	0.4370	0.4238	0.4131	0.3996	0.3890
6	0.5326	0.4819	0.4529	0.4360	0.4239	0.4127	0.4020	0.3932	0.3816
7	0.5049	0.4707	0.4431	0.4271	0.4142	0.4050	0.3959	0.3858	0.3777
8	0.5011	0.4563	0.4344	0.4202	0.4100	0.4010	0.3917	0.3857	0.3804
9	0.4844	0.4478	0.4275	0.4177	0.4089	0.3985	0.3920	0.3862	0.3796
10	0.4725	0.4351	0.4165	0.4066	0.3996	0.3946	0.3889	0.3829	0.3791
11	0.4642	0.4372	0.4179	0.4085	0.4015	0.3944	0.3893	0.3842	0.3798
12	0.4626	0.4358	0.4194	0.4099	0.4033	0.3971	0.3919	0.3878	0.3837
13	0.4576	0.4360	0.4176	0.4091	0.4013	0.3950	0.3910	0.3870	0.3819
14	0.4491	0.4299	0.4153	0.4069	0.4003	0.3955	0.3912	0.3869	0.3843
15	0.4499	0.4284	0.4148	0.4069	0.4017	0.3969	0.3926	0.3881	0.3846
16	0.4472	0.4254	0.4133	0.4065	0.3995	0.3951	0.3912	0.3883	0.3855
17	0.4450	0.4247	0.4144	0.4088	0.4027	0.2976	0.3938	0.3909	0.3880
18	0.4390	0.4233	0.4107	0.4056	0.4012	0.3973	0.3934	0.3904	0.3878
19	0.4400	0.4216	0.4134	0.4074	0.4031	0.3982	0.3940	0.3907	0.3879
20	0.4427	0.4251	0.4148	0.4089	0.4038	0.4005	0.3975	0.3941	0.3912
21	0.4375	0.4220	0.4133	0.4081	0.4034	0.4001	0.3969	0.3943	0.3915
22	0.4377	0.4222	0.4140	0.4080	0.4044	0.4015	0.3977	0.3948	0.3923
23	0.4372	0.4237	0.4157	0.4103	0.4059	0.4024	0.3995	0.3972	0.3943
24	0.4348	0.4218	0.4138	0.4094	0.4058	0.4024	0.3998	0.3974	0.3951
25	0.4346	0.4220	0.4144	0.4099	0.4061	0.4031	0.4004	0.3980	0.3960

(Table computed by S. M. Lee, an improved version of the original table computed by E. Smith.)